Identifying the determinants of housing prices in China using spatial regression and the geographical detector technique

Yang Wang a, b, Shaojian Wang c, *, Guangdong Li d, **, Hongou Zhang a, b, Lixia Jin a, b, Yongxian Su a, b, Kangmin Wu a, b

a Guangdong Open Laboratory of Geospatial Information Technology and Application, Guangzhou Institute of Geography, Guangzhou 510070, China
b Guangdong Academy of Innovation Development, Guangzhou 510070, China
c Guangdong Provincial Key Laboratory of Urbanization and Geo-simulation, School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China
d Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

A R T I C L E   I N F O

Article history:
Received 2 May 2016
Received in revised form 22 October 2016
Accepted 8 December 2016

Keywords:
Housing prices
Spatial regression
Geographical detector
China

A B S T R A C T

This study analyzed the direction and strength of the association between housing prices and their potential determinants in China, from a tripartite perspective that takes into account housing demand, housing supply, and the housing market. A data set made up of county-level housing prices and selected factors was constructed for the year 2014, and spatial regression and geographical detector technique were estimated. The results of the study indicate that the housing prices of Chinese counties are heavily influenced by the administrative level of the county in question. On the basis of results obtained using Moran's I, the study revealed the presence of significant spatial autocorrelation (or spatial agglomeration) in the data. Using spatial regression techniques, the study identifies the positive effect exerted by the proportion of renters, floating population, wage level, the cost of land, the housing market and city service level on housing prices, and the negative influence exerted by living space. The geographical detector technique revealed marked differences in the relative influence, as well as the strength of association, of the seven factors in relation to housing prices. The cost of land had a greater influence on housing prices than other factors. We argue that a better understanding of the determinants of housing prices in China at the county level will help Chinese policymakers to formulate more detailed and geographically specific housing policies.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

With the implementation of its “housing system reform” policy in 1998, China entered a period of rapid growth in housing prices, which have since maintained an annual growth rate of 7% (Lin & Tsai, 2016; Shih, Li, & Qin, 2014). Behind the success of China’s newly developed real estate industry, however, the country faces a serious challenge in the emergence of marked regional differences in housing prices (Li & Gibson, 2013; Zhang, Hui, & Wen, 2015). With the highest national average housing price in May 2014 (in Xicheng District, Beijing) reaching a level 67.7 times that of the lowest (Huzhong District, Da Hinggan Ling, Heilongjiang Province), regional imbalances in housing prices become a burning issue, attracting considerable attention from the nation’s policy makers and scholars alike not least because of their effects on rural–urban migrants’ settlement decisions (Zang, Lv, & Warren, 2015; Li, 2010). The existing literature addressing regional differences in housing prices in China falls into two main categories, which can be differentiated on the basis of the scale of the research (Shih et al., 2014; Wang, Wang, & Wei, 2013). The first category of work comprises provincial studies, which reveal housing prices in provinces in the eastern coastal region to be considerably higher than those in China’s central and western regions. The second category of studies focuses on the city scale, demonstrating the existence of differentiation patterns between spatial agglomerations (between inland areas and the three urban agglomerations of southeast coastal areas) and between urban administrative levels (between provincial capitals and prefecture-level cities) simultaneously. China is a

* Corresponding author.
** Corresponding author.
E-mail addresses: 1987wangshaojian@163.com (S. Wang), ligd@igsnrr.ac.cn (G. Li).

http://dx.doi.org/10.1016/j.apgeog.2016.12.003
0143-6228/© 2016 Elsevier Ltd. All rights reserved.
vast country, and as such it is important to note that regional inequality in housing prices is apparent not only between provinces or cities, but even at the county scale. Taking Harbin as an example, the housing prices of Nangang District, Daoli District, and Daowai District are 8079 yuan/m², 7726 yuan/m², and 7356 yuan/m² respectively, however, the housing prices of Binxian County, Mulan County, and Tonghe County are only 2783 yuan/m², 2952 yuan/m², and 3024 yuan/m² respectively. We note that the housing prices varied significantly across counties in Harbin. If studies of housing prices still focus on provinces and cities, it is impossible to characterise the significant imbalances among counties. There are 2872 counties in China (including the districts of prefecture-level and county-level cities) and county is the basic administrative unit. A county-level analysis would allow researchers to draw more accurate conclusions (Cohen, Ioannides, & Thanapisitkul, 2016), enabling scholars to identify more detailed patterns and mechanisms in the uneven distribution of housing prices.

The factors influencing housing price are complex. An increasing number of studies have been undertaken into these factors in recent years, with the most frequently adopted perspective being one which focuses on supply and demand frameworks (Fortura & Kushner, 1986; Huang & Lu, 2016; Osmadi, Kanal, Hassan, & Sattah, 2015). From the point of view of demand, income and demographic variables constitute the key impact factors (Mankiw & Weil, 1989). Income affects housing demand through the influence it exerts in relation to housing purchasing power (Zhang et al., 2015), a link demonstrated by Nellis and Longbottom's (1981) in their case study of the United Kingdom, which empirically confirmed real income to be the most important factor in housing price. Taking sample cities in Canada as an example, Fortura and Kushner (1986) also found income to constitute a key factor in housing demand, identifying that a 1% increase in household income raises house prices by 1.11%. Similar results were arrived at by Holly, Pesaran, and Yamagata (2010) in their study of the United States and, using unique cross-sectional data on the majority of German counties and cities for 2005, Bischoff (2012) also identified an interdependence between real estate prices and income. Demographic variables are a second important set of factors influencing housing demand (Buckley & Ermisch, 1983; Hui, Wang, & Jia, 2016). Just as population growth and population shrinkage will respectively increase or decrease housing demand, thereby affecting housing price (Capozza & Schwann, 1989; Maennig & Dust, 2008), low population mobility will also result in decreases in housing demand (Gabriel & Nothaft, 2001). In addition, urban economic fundamentals—i.e., indicators such as per capita disposable income, population, the unemployment rate, and housing vacancy rates—have also been found to affect housing prices (Shen & Liu, 2004).

Adopting a housing supply perspective (Davidoff, 2013), Glaser and Gottlieb (2009) hold that housing prices represent the interaction of supply conditions. Land supply and housing construction costs, for instance, can both directly affect housing supply and influence housing prices (Li, 2010)—while Potepan (1996) found land supply to be the most important factor affecting housing supply, Holmes, Otero, and Panagiotidis (2011) have showed that opportunity costs for builders (when considered in relation to alternative forms of investment), as well as construction costs, both influence housing prices. Bischoff (2012) found land price to constitute a key factor affecting housing supply and housing prices. Further, housing supply elasticity must also be taken into account, as this limits affordability for buyers (Glaser, Gyourko, & Saks, 2006; Quigley & Swezoda, 2010).

Housing prices have thus been studied from both supply and demand perspectives, and a diverse range of factors affecting urban housing price have been identified. Combining these two perspectives, Manning (1986) developed an equilibrium model to explain inter-city variation in housing price appreciation. Manning’s empirical model, which considered 16 independent variables reflecting both housing demand and housing supply, was able to account for 68.8% of housing price appreciation. Based on a supply and demand framework, Malpeazzi (1996) similarly analyzed inequality in housing prices between cities by looking at income, population mobility, non-economic factors (for instance, topography), and the management/legal environment. Using panel data for 62 metropolitan in the United States, Capozza, Hendershott, Mack, and Mayer (2002) found that city size, real income growth, population growth, and real construction costs all correlate positively with housing price. In addition, Holmes et al. (2011) found labor and capital mobility to constitute key factors affecting housing prices. In addition to supply and demand frameworks, the housing market itself also constitutes an important determinant in relation to housing prices and an important focus in existing literature (Mahalik & Mallick, 2016; Pillaiyan, 2015). Adopting such a perspective, Hwag and Quigley (2006) found vacancies and residual construction activity in the owner-occupied housing market are also an important factor affecting housing price. From this brief review of existing literature, it is apparent that research into housing prices would benefit from adopting a comprehensive framework able to take into account all three perspectives of housing supply, housing demand, and the housing market.

Importantly, the majority of the studies mentioned above did not take into account spatial effects, including spatial spillovers, spatial dependence, and the spatial heterogeneity of housing prices among neighboring regions (Canarella, Miller, & Pollard, 2012; Bitter, Mulligan, & Dall’erba, 2007). It is almost inevitable that spatial autocorrelation or spillovers effects exist in geographic data (Chiang & Tsai, 2016; Yu, 2015). Cohen et al. (2016) examined the spatial effects in house price dynamics and found spatial diffusion patterns in the growth rates of urban house prices from 363 metropolitan statistical areas in the United States for 1996 to 2013. In fact, such autocorrelation, if ignored, can lead to biased or even misleading conclusions (Ma & Liu, 2015). Kuete and Pede (2011) explicitly incorporated locational spillovers through a spatial econometric adaptation of a vector autoregression model. Their results suggested that the inclusion of spatial information led to significantly lower mean squared forecast errors. This is particularly pertinent in the study of housing prices, given that existing literature has clearly identified that marginal prices vary across space, and it constitutes a major oversight not to take this into account. Analyses of inequality in housing prices should therefore, in addition to the consideration of housing supply, housing demand, and the housing market, also address the influence exerted by spatial effects.

Through their study on the impact factors influencing housing prices in China—a subject particularly pertinent to the present study—Chen, Guo, and Wu (2011) found rural-urban migration and urbanization to have had important impacts in relation to provincial housing prices. Working in a similar vein, Li and Chand (2013) found levels of income, construction costs, impending marriages, user costs, and land prices to constitute the primary determinants of house prices in China’s 29 provinces. From a sample made up of 14 Chinese cities, Shen and Liu (2004) found economic fundamentals and variables relating to urban households (i.e., per capita disposable income, population, unemployment rate, and vacancy rate) to constitute key factors influencing housing prices. Wen and Goodman (2013) found urban land price to maintain an endogenous interrelationship with housing price in 21 Chinese provincial capitals. Similar results were also arrived at by Du, Ma, and An (2011) in relation to four major cities in China. Wang and Zhang...
(2014) found that the urban *hukou* population,\(^1\) wage income, urban land supply, and construction costs are important factors in the analysis of inequality in urban housing prices. In addition, we note that spatial spillover effects (in particular, the diffusion effect between core and periphery) have also been found to exist in relation to housing price in China (Shih, Li, & Qin, 2014).

While this previous research has certainly enriched our understanding of the main impact factors determining housing prices in China, a number of shortcomings are also evident in these previous studies. Importantly, the existing research has to a large extent focused on the direction of factors, but neglected the magnitude of those factors. Both the direction and magnitude of factors should, as it is argued, be examined. In addition, most Chinese studies have either focused on the level of provinces, or else on the city level, to the exclusion of the county level. Building on this previous research, the present study firstly adopted a tripartite perspective, addressing housing supply, housing demand, and the housing market. It then employed spatial regression models and the geographical detector technique in order to examine the direction and magnitude of factors affecting housing prices. In addition, cross-sectional data for China’s 2760 counties (excluding 112 counties for which data was unavailable) was used in order to provide further accuracy to the interpretation of those factors.

The paper proceeds as follows. Section 2 focuses on data and methodology, presenting the conceptual framework of the study. Section 3 presents a discussion of the results of the study, addressing the spatial pattern of housing prices, identifying the factors behind housing prices, and examining the direction and magnitude of those factors. Section 4 sets out the main conclusions.

2. Material and methodology

2.1. Conceptual methodology

In line with our aim to identify the determinants of housing prices in China at the county level, we designed a conceptual framework that would allow us to address the spatial complexity of China’s housing prices and synthesize the multiple driving forces at work (Fig. 1). While we recognize that the regional inequality present in Chinese housing prices is sensitive to spatial scales of analysis—it can be analyzed at the provincial, regional, city and county levels—we advocate that regional differences are most evident at the county level, rather than the provincial or city level. As such, the present study adopted a finer scale of analysis than that of previous studies, considering housing prices at the county level. The analysis procedure was as follows. Using data for 2760 counties in China, the study aimed to identify the influencing mechanisms in relation to housing price, from the point of view of housing supply, housing demand, and the housing market. As such, we began by selecting a series of impact factors able to address this from a tripartite perspective. Spatial regression techniques were then employed in order to determine the significances and directions of the impact factors. We then investigated the magnitude of the significant factors in greater detail using the geographical detector technique. Finally, interpretation of the result was conducted in order to explain the manner in which the factors influenced housing prices.

2.2. Initial selection of factors behind housing prices

Building on previous studies and in line with both the housing supply-demand perspective and the housing market perspective, we began by selecting a number of exploratory variables in relation to housing prices. From the perspective of housing demand, the following variables were selected: the proportion of renters (%), per capita living space (m\(^2\)/person), the urban population of the floating population\(^2\) (%), and the average wage of urban employees (Yuan). From the perspective of housing supply, land prices (10\(^4\)Yuan/ha) was selected as the basic indicator. In support of a housing market perspective, we selected the proportion of the population working in the real estate industry (%) and the proportion of employment occurring within the tertiary sector (%) to use as basic indicators.

Table 1 provides a statistical description of the variables, as well as the expected directions of variables for housing price, based on the results of previous studies. Fig. 2 displays a scatter plot and distribution overlay of housing price and its determinants in the form of a box plot, where the bottom and top of the box represent the 25th and 75th percentile.

The reasons of variables selection are as follows.

The proportion of renters. On the one hand, the number of renters represents the size of urban housing demand (Goodman, 2003). On the other hand, the larger the proportion of renters in a given city, the greater is the difficulty of urban housing purchases and the higher are housing prices. Therefore, the proportion of renters represents the rigid demand of citizens for housing and reflects the difficulty of housing purchases.

Living space (denoted by per capita living space). Taking China as an example, Zhang (2015) found that the smaller the living space, the worse is the access to the house. Therefore, living space represents rigid and improved demand for housing. From the perspective of housing demand, theoretically, living space is negatively correlated with housing prices.

Floating population (denoted by the urban proportion of the floating population). Previous studies have found correlations between the floating population and housing prices (Chen et al., 2011; Gabriel, Shack-Marquez, & Wascher, 1992). Some scholars have even found that the floating population is positively correlated with housing demand (Gabriel & Nothaft, 2001). In addition, the limited housing access of the floating population may raise housing demand (Jiang, 2006). Therefore, the higher the proportion of the floating population in a city, the higher are prices.

Wage level (denoted by the average wage of urban employees). A large number of previous studies have found that wage level is positively correlated with housing prices (Antolin & Bover, 1997; Hoehn, Berger, & Blomquist, 1987). From the perspective of housing demand, wage level is representative of income, which is recognized to have a positive relation with housing prices (Gallin, 2006). From the perspective of housing supply, Van Nieuwerburgh and Weill (2010) found that an increase in wages reduces local housing supply and further raises housing prices. They also found that when wage inequality increases in U.S. metropolitan areas, the imbalance in housing prices increases accordingly.

Cost of land (denoted by land prices). Housing prices and land prices have an endogenous relationship (Wen & Goodman, 2013). From the perspective of housing supply, a shortage in land supply

---

\(^1\) *Hukou* is the Chinese household registration system. Each person has a hukou (registration status). In Chinese cities, some people (usually rural migrant workers) are prevented from being registered as local hukou. These are called ‘non-hukou migrants’ (or the floating population). Hence, the urban population consists of the hukou population and non-hukou migrants. The hukou population receives basic welfare and government-provided services (e.g. access to local schools, urban pension plans), whereas non-hukou migrants are not eligible for these (Chen, 2010).

\(^2\) Floating population is the person without local hukou (also called ‘non-hukou migrants’). The population of a Chinese administrative region unit is composed of hukou population and floating population. Floating population is not eligible for regular urban welfare benefits and other rights that are available to those with urban hukou (Chen, 2010).
increases land prices (Zabel & Dalton, 2011), and such a rise in land prices raises housing prices (Hui, 2004). From the viewpoint of cost, the cost of land is the most important factor behind housing costs (Wen & Goodman, 2013). Therefore, land prices are positively correlated with housing prices.

**Housing market** (denoted by the proportion of the population working in the real estate industry). Housing prices are responsive to housing market conditions (Zhu, 2006). Hwang and Quigley (2006) found that employment in local housing markets is an important determining factor of housing prices in U.S. metropolitan regions. The proportion of the population working in the real estate industry can reflect the activeness of the housing market, which is positively correlated with housing prices.

**City service level** (denoted by the proportion of employment in the tertiary sector). Roback (1982) found that the level of local amenities influences where residents live. City service level is thus the core element of a city’s attractiveness and competitiveness. Therefore, we use the proportion of employment in the tertiary sector to proxy for city service level.

### 2.3. Data

County-level housing price data for 2014 were obtained from three websites: Xitai data (禧泰数据), which is the China’s biggest real estate transaction database, accessed from the website http://www.cityre.cn/cityCenter.html; Haowu data (好屋房价), from the website data platform http://jia.haowu123.com/; and 58Tongcheng data (58同城) were collected from the website http://*.58.com/ through artificial data mining. All the housing data was acquired for the month of June, in 2014.

The seven socioeconomic indicators used in this study were derived as follows: POR, PCLS, UPOFP, POPWIREI, and POEOWTS were derived from the tabulation of data taken from the Population Census of the People’s Republic of China by county; LP was derived from the China Statistical Yearbook of Land Resources; and AWOUE was derived from the China Statistical Yearbook for the Regional Economy, China County Statistical Yearbook, and China City Statistical Yearbook. Data were collected for 2010. On the one hand, the data in 2010 are more comprehensive than any other years because of the Sixth National Population Census of China (which occurs every 10 years). On the other hand, according to existing research, housing demand, housing purchasing power, and land price factors have a three-year lagged impact on housing prices (Chen & Yang, 2013). Therefore, it is feasible to collect socioeconomic indicators from 2010.

### 2.4. Methods

#### 2.4.1. Spatial regression models

In order to decipher the sources of regional inequality in housing prices, it is necessary to first understand the mechanisms behind the differentiation of housing prices. Studies of regional inequality often employ a variety of conventional regression measurements, such as ordinary least squares (OLS) (most commonly used), generalized method of moments (GMM), and maximum likelihood (ML). With the development of spatial data analysis—undertaken through the spatial error model (SEM) or spatial lag model (SLM), amongst other methods—regressions are able to explicitly take into account spatial effects. This research used both sets of complementary regressions in order to investigate the directions of impact factors affecting housing prices at the county level in China. A spatial lag model (SLM) was used to reflect the impact of spatial units on other near units in the whole region. The
SLM model can be specified as follows:

\[
y_s = \hat{\beta}_0 + \hat{\beta}_1 x_{si} + \mu_s + \varepsilon_s, \quad \varepsilon_s \sim i.i.d \left(0, \sigma^2 \right)
\]

(1)

where \( s = 1, \ldots, 2760 \) denotes the spatial location of counties in mainland China; \( y_s \) is the dependent variable (housing price); \( x_{si} \) is the seven independent variables, including PER, PCLS, UOFP, AWOUE, LP, POPWIREI, and POEOWTS; and \( \beta \) is the local regression parameters to be estimated. \( W \) is a diagonal weighting matrix.

In the model setting, its variables related to the dependent variable (i.e., variables with concealment or that are unable to be accurately quantified) can be inadvertently omitted. Spatial autocorrelation may exist among the variables. Given that the independent error term may impact on the spatial spillover effects that exist between geographic units, spatial autocorrelation with no independent error term may lead to biased or even misleading conclusions being drawn. The spatial error model (SEM) can solve this problem of the error term. The SEM is generally based on the following autoregressive model:

\[
y_s = \hat{\beta}_0 + \hat{\beta}_1 x_{si} + \mu_s + \varphi_s, \quad \varepsilon_s \sim i.i.d \left(0, \sigma^2 \right)
\]

(2)

where \( \rho \) is the spatial autocorrelation coefficient of error term; \( \varphi \) is the error term of spatial autocorrelation.

2.4.2. The geographical detector technique

The presence of several significant determinants in relation to housing price necessitated the use of the geographical detector technique, a method proposed by Wang et al. (2010) which has been widely used in analyzing the effects of underlying factors on local disease risk. The geographical detector method does not require any assumptions or restrictions to be made with respect to explanatory and response variables. Despite the wide usage of the geographical detector technique in the fields of public health, the use of geographical detectors in the study of the mechanism of housing price is to date limited. In this research, we assumed that if housing price was influenced by a particular factor, the distribution of this factor would be similar to that of housing price in geographical space—in other words, if a particular factor was found to affect, then the spatial distribution of this factor and the distribution of housing price would be consistent (Hu, Wang, Li, Ren, & Zhu, 2011). The power of determinant \( D \) to the housing price effect \( H \) can be written as:

\[
P_{DH} = 1 - \frac{1}{nDH} \sum_{i=1}^{m} nH_i \sigma^2_{DH_i},
\]

(3)

where \( P_{DH} \) is the Power of Determinant of factor \( D \); \( m \) denotes the number of the sub-areas; \( n \) denotes the number of counties in the areas of the study area; \( nD_i \) denotes the number of counties in sub-areas; \( \sigma^2_{DH} \) is the global variance of housing prices in the areas of the study area; \( \sigma^2_{DH_i} \) is the variance of housing prices in the sub-areas; usually the value of \( P_{DH} \) lies in \([0, 1]\). The larger the value of \( P_{DH} \), the greater is the impact of \( D \) on the housing prices.

3. Results and discussion

3.1. The spatial pattern of housing prices

County-specific housing prices are shown in Fig. 3; these range from 963 Yuan/m² to 65,227 Yuan/m² with a median of 3731 Yuan/m² and a mean of 4882.84 Yuan/m². Housing prices were found to vary greatly between the northeastern counties and their southwestern counterparts. Counties with high housing prices were mainly located in the southeastern-coastal and Beijing-Tianjin areas. The average housing prices of the eastern region were remarkably higher than those of the western and central regions. We also found that the higher administrative level of a given county, the higher its housing prices.

Further, the housing prices of urban areas of provincial capital cities were always higher than those of prefecture-level and county-level cities. Descriptive cluster statistics for housing prices are presented in Fig. 4.

We then calculated global Moran’s \( I \), an index which measures spatial autocorrelation or spatial agglomeration. The global Moran’s \( I \) of the county-level housing prices was found to be 0.3538 (significant at 95 percent confidence level via the randomization assumption), a result which indicates a relatively strong, positive spatial correlation. We used the local Moran’s \( I \) in order to analyze the local clusters of housing prices (Fig. 5), revealing that the high-high clusters comprised of: Beijing-Tianjin (Region 1), the Yangtze River Delta (Region 2), the central part of the Fujian area (Region 3), and Guangzhou-Shenzhen (Region 4). The low-low clusters were, in comparison, found to be located in the border areas of Henan-Shanxi-Shaanxi-Hubei province (Region 5) and the border region of Hunan and Guizhou provinces (Region 6). The low-high outliers were found to be located within the ring of the capital economic circle (Region 7).

3.2. Factors influencing housing prices

To understand the significant inequality evident in the housing prices of China’s counties, it is necessary to study the influencing mechanisms behind those prices. Regression remains the most
favored method for exploring the impact factors of housing prices at various scales in existing literature on the topic. OLS regression analysis, with easily defined dependent and independent variables, is frequently used to estimate the underlying determinants. However, OLS regression would not be able to account for spatial interaction or spillover effects between study units. In comparison to OLS regression, spatial regression (both lag and error models) can reveal spatial autocorrelation. In line with the hypothesis that “while everything is related to everything else, near things are more related than distant things (Tobler, 1970),” spatial regression can demonstrate whether variables present in proximate areas (in this study, proximate counties) are more important than those present in distant areas. Given these qualities, we undertook spatial regression analysis in order to investigate potential influence factors. For comparison purposes, OLS regression was also carried out. The results of the regression analyses (OLS and spatial lag and error models) are reported in Table 2.

Goodness-of-fit statistics such as the AICs and log-likelihoods can be used to estimate the fitting degree of regressions. AICs and log-likelihoods for the present study are set out in Table 2; these indicate that the data were better fit using spatial analysis techniques than an OLS regression. The AIC for the OLS model was found to be 51,741.34, while for the spatial error and lag models they were 50,031.59 and 48,671.56 respectively. Ignoring potential spatial effects in the regression analysis would, as these results reveal, reduce the model’s effectiveness (Table 2). In addition, robust Lagrange multiplier tests were undertaken, the results of which indicated that the spatial lag model was more suitable than the spatial error model for the purposes of this study.

The results of the spatial lag model revealed that POR, UPOFP, AWOUE, LP, POREI, WP, and POEWOTS positively and significantly influence housing price at the county level in China. Conversely, PCLS was identified as having an inhibitory effect on housing price. After identifying the significances and directions of the effects of the potential factors on housing prices, we then proceeded to estimate the magnitude of the impact of those factors.

Each factor was then allocated to one of five sub-regions in geographical space according to the original value: it was thus determined to exhibit either a high level (10%), a high-middle level (20%), a middle level (40%), a low-middle level (20%), or a low level (10%). Based on this rule, we identified the threshold of the sub-regions of seven detectors; their distributions are displayed in Fig. 5. Using the geographical detector technique, we subsequently estimated the magnitude of the impact of these seven factors (Fig. 6). As shown in Fig. 6, the factor detector discloses the influence of the seven factors on housing prices, which can be ranked by the power of determinant (PD) value as follows: LP (0.383688) > POR (0.329070) > POPWIREI (0.31374) > AWOUE (0.281561) > POEWOTS (0.220390) > UPOFP (0.156177) > PCLS (0.007615). This result shows that land prices can predominantly explain the spatial variability of housing prices, followed by the proportion of renters and the proportion of the population working in the real estate industry, while per capita living space was found to have a weak influence.

Taken together, the results of the spatial regression and the geographical detector technique reveal a great deal about the mechanisms underlying the uneven housing prices of China’s counties. Firstly, the cost of land was found to have exerted strong influence in relation to housing prices in China. The results indicate that high cost of land is the most important factor in raising house prices. Fig. 6a shows that the distribution of this factor is consistent with that of housing prices in geographical space. As we all know, the cost of land is the foundation of housing prices. According to the study undertaken by Zeng and Zhang (2013), the proportion of China’s urban land price of housing price rose from 9.0% in 1998 to 24.3% in 2011. Land cost inequality has become an important factor in residential cost differences, and changes in the price of land cause changes in the price of housing. As discussed by DiPasquale and Wheaton (1996), raising land prices will decrease the supply of land—if plot ratio is constant, the supply of housing will decline; if demand is constant, housing prices will increase. Thus, from the perspective of housing supply, the cost of land can be said to positively correlate with housing price.

Second, the spatial regression results revealed that the influence of the proportion of renters was significant, as it was found to have a higher PD (0.329070). While Chinese culture incorporates a strong concept of home ownership, renting is considered a transitional solution. In China, renters are the largest group driving housing demand; as this study reveals, the higher the proportion of renters, the higher housing prices. For instance, on one hand, in Shanghai’s Huangpu district (which is located in the center of the city), the proportion of renters was 66.82% and the housing price was 45,172 Yuan/m²; on the other hand, in Luyi county in Hebei province, the proportion of renters was only 0.01% and housing prices only 2709 Yuan/m². Exceptions exist to this general rule—for instance, while the counties with rich oil and gas resources in northwestern China possess the highest proportion of renters (about 10%–20%), housing prices are relatively low (Fig. 6b). This is because a large proportion of short-term migration population is present in these counties. Such floating populations have relative low purchase intentions as a result of the shorter term of their inhabitation.

Third, the housing market—measured in this study by the proportion of the population working in the real estate industry—can directly reflect the investment demand for housing and the maturity of the real estate market. As our results show (see Table 2 and Fig. 7), the housing market activeness positively correlates with housing prices, via a PD value of 0.321374. These results are consistent with Zhu (2006) study, which found housing prices to be associated with environmental changes in regional housing markets, implying that transactional friction in real estate markets will impact on housing prices and further that market turnover is positively related to the housing prices (a finding supported by the work of Caplin and Leahy (2011)). The results shown at Fig. 6c indicate that the higher the city administrative level of a given county, the higher its housing prices. For instance, the proportion of the population working in the real estate industry in the Changning and Jing’an districts in Shanghai is over 5%, places where housing prices are up to 41,820 and 52,220 Yuan/m² respectively. Conversely, in the 978 counties, mainly located in inland China, with the proportion of the population working in the real estate industry less than 0.1%, the average housing price was only 3145 Yuan/m².

Fourth, the level of wages contributes a rather prominent, positive influence in relation to housing prices in China, generating a PD value of 0.281561. While wage level can reflect the affordability of housing for local buyers, it is also the foundational factor in determining whether the housing prices continue to rise or not. As a result, wage level exerts a strong impact on housing demand. Regional wage levels are, moreover, one of the important factors determining an attractive labor market, a characteristic which may determine regional competitiveness. If wages are not competitive, regional labor supply will gradually decrease, thereby reducing the demand for housing (DiPasquale & Wheaton, 1996). In China, significant inequality in wages exists among the country’s regions—for instance, while average wages in the urban areas of Beijing and Shanghai are more than 80,000 Yuan/year, 184 counties exist which maintain an average wage that is less than 20,000 Yuan/year (Fig. 6d). Regional socioeconomic inequality is thus mainly characterized by wage differences, and wage inequality feeds back into
housing prices through supply and demand mechanisms. Through this feedback mechanism, wage level constitutes an important factor in determining housing prices in China.

Fifth, city service level is measured as the proportion of the working population engaged in employment in the tertiary industries. The results of this study show that city service level significantly affects rising regional inequalities of housing prices, generating a PD value of 0.220390. The rapid development of the service industry drives the growth of the urban economy and further promotes a rapid increase in consumption. In other words, optimizing the industrial structure and the consumption structure will increase a county’s attractiveness in terms of regional employment. High levels of employment in the service industries result in increasing housing demand, which is directly reflected in higher housing prices. For instance, the high level of service industry development in Guangzhou corresponds with a housing price which is double that of Dongguan and Foshan in Guangdong province. However, there exist exceptions to this broad trend: in the north of Inner Mongolia and Xinjiang provinces, while the tertiary industry employment proportion is high, housing prices remain low (Fig. 6e).

Sixth, the percentage of the population which is floating contributes a small, positive influence on housing prices in China, through a PD value of 0.156177. The floating population plays an important role in regional labor markets, and is also an important reflection of an area’s attractiveness. Young people tend to migrate between areas, and as population inflows increase, housing demand increases. In eastern and central China, the spatial distribution of population inflows and the distribution of housing price were found to be consistent (Fig. 6f). In western China, regions with rich oil and gas resources were found to maintain large floating populations, despite displaying low housing prices—this may account for the relatively low PD value.

Last, per capita living space was found to contribute a small, negative influence on housing prices in China—its PD value was only 0.007615. Theoretically, improvements in urban residential environments decrease in line with growth in per capita living space, rendering housing price growth static. However, in contemporary China, housing growth still mainly comes from the rigid demand from potential purchasers. Thus, living space exerts little impact in relation to housing prices. In addition, from Fig. 6g, we found that regions with high per capita living space were mainly located in the Yangtze River Delta, a distribution which maintains an insignificant negative relationship with that of housing prices. Per capita living space therefore can be concluded to exert a minor negative impact in relation to housing prices.
4. Conclusions

This study analyzed the distribution of and driving forces behind housing prices in China. We constructed a county-level data set for housing prices and their determinants at the county-level, employing regression models and the geographical detector technique in order to identify the magnitude of the impact exerted by significant factors on housing prices. The results are summarized as follows.

The pattern of China’s housing prices was found to be characterized by administrative level and spatial dependence. On the one hand, this means that housing prices positively correlate with the administrative level of counties. The higher the administrative level of a given city, the higher are housing prices. On the other hand, the spatial differentiation of housing prices is mainly between urban agglomerations of coastal areas and inland areas. Specifically, counties with a higher administrative level maintain higher housing prices than those which exist at a lower administrative level. In addition, the housing prices of coastal counties are also higher than those of inland counties. When considering the core-periphery structure of regional development in China, it is worth noting that counties in the core area have benefited more from economic growth in terms of their housing prices than counties in the periphery. In addition, on the basis of results obtained using Moran’s I (global and local), the study revealed the existence of significant spatial autocorrelation (or spatial agglomeration) in the modeled data. We found that the high-high clusters were Beijing-Tianjin and the Yangtze River Delta, etc., and the low-low clusters were the border areas of Henan-Shanxi-Shaanxi-Hubei province and the border region of Hunan and Guizhou province. Compared to previous studies which considered 31 provinces (Chen et al., 2011; Zang et al., 2015) or 35 major cities (Taltavull, Kuang, & Li, 2012), by considering the county level this study was able to identify a more detailed spatial pattern in relation to China’s housing prices.

The factors underlying the uneven housing prices seen in Chinese counties can be revealed through regression analysis. We

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Estimation results of regressions.</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS model</td>
<td>Estimate</td>
</tr>
<tr>
<td>(Intercept)</td>
<td>−227.6737</td>
</tr>
<tr>
<td>POR</td>
<td>24.60935</td>
</tr>
<tr>
<td>UPOF</td>
<td>0.0007130805</td>
</tr>
<tr>
<td>AWOUE</td>
<td>0.00039646</td>
</tr>
<tr>
<td>LP</td>
<td>2.765854</td>
</tr>
<tr>
<td>POPWIREI</td>
<td>1457.208</td>
</tr>
</tbody>
</table>

Adjusted R-squared: 0.733995, F-statistic: 1084.81; p-value: 0.0000000

Spatial error model

| (Intercept) | 1479.709 | 292.9998 | 5.050206 | 0.0000004 |
| POR | 20.99774 | 5.75936 | 3.6373468 | 0.0001676 |
| UPOF | 0.000538408 | 0.0001407248 | 3.825963 | 0.0001303 |
| AWOUE | 0.0384822 | 0.005113064 | 7.52625 | 0.0000000 |
| LP | 1.803015 | 0.06906678 | 26.10539 | 0.0000000 |
| POPWIREI | 815.6232 | 68.60551 | 11.8886 | 0.0000000 |
| POEOWTS | 13.28594 | 4.30163 | 3.873268 | 0.0000000 |

Adjusted R-squared: 0.796388; Log likelihood: −24466.3 for error model; AIC: 50031.59, (AIC for OLS: 51741.34)

Robust Lagrange multiplier test: 1792.331 on 1 DF, p-value: 0.0000000

Spatial lag model

| (Intercept) | −654.8888 | 182.2316 | 3.683011 | 0.004959 |
| POR | 8.222733 | 4.891887 | 1.680892 | 0.0927839 |
| UPOF | 0.0005393984 | 0.0001845976 | 3.415289 | 0.0000000 |
| AWOUE | 0.04147592 | 0.00483845 | 8.572151 | 0.0000000 |
| LP | 1.306428 | 0.0687123 | 22.97168 | 0.0000000 |
| POPWIREI | 699.9601 | 68.40823 | 10.23211 | 0.0000000 |
| POEOWTS | 9.124655 | 3.17219 | 2.877344 | 0.0040106 |

Adjusted R-squared: 0.873652; Log likelihood: −24327.752987 for lag model; AIC: 48671.56, (AIC for OLS: 50741.0)

Robust Lagrange multiplier test: 2069.52 on 1 DF, p-value: 0.0000000

Fig. 5. Local Moran’s I for China’s housing prices at the county level.
Fig. 6. Maps of seven factors of housing prices in China.
found that the proportion of renters, the floating population, the wage level, the cost of land, the housing market, and the city service ability all significantly positively influenced housing prices, while living space exerted a significant negative influence. Importantly, the data was found to maintain a better fit using spatial analysis techniques than conventional regression measurements, supporting the potential of spatial regression to reveal a spatial dependence relationship in housing prices. Our results indicate that regional housing development is bounded by local conditions, and that spatial analysis techniques can in fact capture the effects driving housing prices.

We used the geographical detector technique to analyze the strength of the association between the factors studied and housing prices in China. Our results indicate that the cost of land makes the greatest contributions to housing prices. This result also suggests that housing supply is the dominant influencing factor behind China’s housing prices in contrast to studies from the perspective of housing demand. It is followed by the proportion of renters and the state of the housing market. These two factors have seldom been adopted in previous studies; however, we found that they were the core factors driving housing prices at the county level in the present study. Wage level, city service ability, and the floating population were also identified as general influencing factors, the strength of which was found to be lower than the core factors. While living space did exert significant influence on housing prices, its impact was relative small compared to other factors analyzed.

From a methodological perspective, this study highlights the promise of spatial analysis techniques and the geographical detector techniques in understanding the factors influencing housing prices. These methods hold great potential for the generation of further knowledge around these issues through future studies. Our empirical analysis explored the main factors in housing price differentiation at a more detailed scale than existing studies (county-level), providing a reference for the government to formulate more detailed and geographically specific housing policy.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (41601151, 41401164 and 41501175) and the Natural Science Foundation of Guangdong Province (2016A030310149).

References


Fig. 7. The power of determinant for the 7 factors guiding the housing price effect.


